An Explanation Method of Unfamiliar Tourist Spots based on Roles of User's Familiar Spots

Michael Shell, *Member, IAENG,* John Doe, *Senior Member, IAENG,* and Jane Doe, *Fellow, IAENG*

***Abstract*—In recent years, when planning tourist spots, planning is often made by utilizing tourist information on the Web. However, after deciding the area you want to visit from many areas, the user also needs enormous amount of time and effort to find tourist spots that match your image. In addition, there are cases where the user feels expectation and anxiety with respect to the unvisited spot. In this research, in order to support understanding of users' unknown spots, we propose explanatory method that supports the understanding by fitting the features of tourist spots that have already visited to unvisited spots. In order to emphasize the features of the tourist spots themselves, extraction of features of each tourist spot is done by work using all reviews of tourist spots entered by the user, all reviews of tourist spots in the target area. We also conduct an evaluation experiment to construct the prototype system and verify the effect of the explanatory information between the visited spot and the unvisited spot.**

***Index Terms*—tourist spots, understanding support, reviews, cosine similarity, TFIDF, harmonic mean**

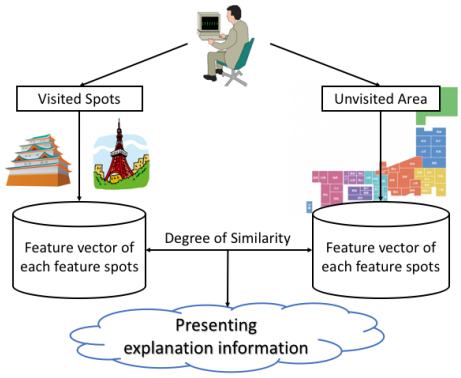

Fig. 1. An explanation method of unfamiliar tourist spots based on roles of user's familiar spots

I. Introduction

WHEN deciding the travel destination, the traveler selects tourist spots by planning a travel plan, watching tourist spots search sites and books related to tourist information. However, after deciding the area you want to visit from many areas,and further from their many tourist spots in the area is not easy to find. In the case where the tourist spots desired to go are not decided, it is considered that it is more likely to decide tourist spots by looking at ranking and recommendation information. At this time, the image for the tourist spots selected by the user becomes ambiguous, which may cause anxiety.

In recent years, the speed of development of tourism industry and social networking service is accelerating, and the number of users who post reviews on tourist spots experienced to the tourist spot search site is increasing. In order to effectively understand various tourist spots, it is essential to consider the correspondence between unknown information and existing information based on existing information. This way of thinking is equivalent to analogy which applies to the things by previous experiences (called the base), or problems (called the target). For example, whereas unknown spots such as "Kanazawa's Nisityayagai", if you explain that it is similar to the already visited "Kyoto Hanamikoj", it may make it easier to understand the image.

Manuscript received April XX, 20XX; revised June XX, 20XX. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine.

M. Shell is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA e-mail: (see http://www.michaelshell.org/contact.html).

J. Doe and J. Doe are with Anonymous University.

In this research, in order to support understanding of users' unknown spots, we propose explanatory method that supports the understanding by fitting the features of tourist spots that have already visited to unvisited spots. Specifically, from the already visited spot and the unvisited area entered by the user, we use the review to extract the unique features of each spot in the already visited spot and the unique features of each spot in the unvisited area, compare and present explanatory information. With this prototype system, users aim to support understanding of unvisited spots. Fig. 1 is a conceptual diagram of the proposed method.

The structure of this paper is as follows. Section II describes related research. Section III gives an overview of the proposed method. Section IV describes evaluation experiments and considerations to verify the effect of the constructed prototype system. Section V describes with conclusions and future work.

II. Related Work

A. Research using user history

Many researches on retrieval and recommendation system using the user's experience history have been published. Kurashima et al.[1] proposed a travel route recommendation method using geotag information of photos posted to Flickr as a travel history of people. In this method, it is assumed that it is easy to move from a user's present location to a place easily accessible to the user's interest, and a behavior model is generated. That geotagged photo aggregation of users can be regarded as personal travel history when sorted by time information, and we generate user behavior model using geotag information. Kitamura et al.[2] proposed a method of recommending sightseeing spots based on estimating user's preferences of travel plan from past personal travel photographs using general object recognition. Using an object

recognition system to acquire keywords of subject information taken in the photos and represented the co-occurrence of the keywords by a graph visualization technique. In addition, present a user interface that visualizes a graph with travel photos based on our graph visualization technique. Cheng et al.[3] used photographs of freely available community contributions to focus on personalized trip recommendations, considering specific user profiles or attributes, suggested to consider personalized travel recommendations.

B. Research on analogy

Analogies were pointed out as contributing to creative thinking[4]. Analogical thinking works when acquiring a concept (called the target) from known knowledge (called the bases)[5]. Many of the researches on analogy are given the base learning data and targeted problems, and the problems are solved by mapping the features of things to the characteristics of the problem[6]. Gick et al. designed to investigate the use of analogies between disparate domains as a guide to finding solutions for an ill-defined problem. Some studied about how to give learning data and functions[7], and clarified whether to solve the problem depending on the degree of cognitive proficiency[8]. In many of the conventional research including these, after giving bases and targets for analogy, we solve problems according to a certain procedure. There are three types of structural similarities"similarity of object level" determined by the number of shared features, "relationship similarity" based on the degree of sharedness of relationships existing in the base and the relationship existing in the base, and similarity in the title solution or target level There is a certain "pragmatic similarity"[5], [9].

In the conventional method of using the user's experience history, many researches that analyze the geotag information of the history photograph and make it user's preference are performed. In addition, it is well used in learning support on analogy technology. In this research, using the review of visited spots and unvisited areas, the relative features of each spot in each set of user visited spot set and unvisited spot set are determined and associated, thereby supporting understanding of spots explanation information can be presented. Moreover, because the quality of analogy is treated explicitly, it is considered to be similar to the similarity of structure "similarity of relationship level" by this research.

III. An Explaination Method of Unfamiliar Tourist Spots

We propose an explanation method of unfamiliar tourist spots based on roles of user ' s familiar spots Specifically, first, the user inputs a plurality of tourist spots that have been visited and tourist spot area information that user wishes to visit. Use the already visited spot review vector to find the feature vector for each visited spot. Similarly, the feature vector of each spot in the area is obtained for an unvisited spot. Next, we associate unvisited tourist spots with features similar to the difference features between the visited spot review vector and the unvisited spot review vector. Finally, explanation method for supporting understanding of unvisited spot is defined using TFIDF and presented to the user.

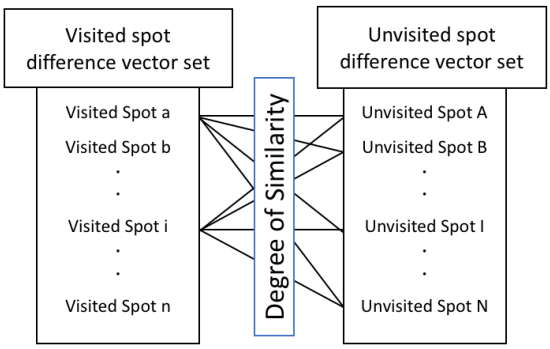

Fig. 2. Concept of similarity calculation

A. Generating feature vector from reviews of spot

Review vectors of previously visited spots and unvisited spots are created using a discriminated (original) review with the morphological analyzer[10] "mecab-ipadic-NEologd"[1]. After that, using Distributed[2] Bag-of-Words of Doc2Vec, we use a vector created in 300 dimensions using all reviews of each spot. In this paper, we will use the review data obtained from "Jalan"[3] until the end of September 2016.

B. Relative features for role of tourist spots

In this research, features of tourist spots make use of relative features. The relative feature is a unique feature when a specific tourist spot is compared with other tourist spots included in a set of tourist spots. As an example, consider the case where there are "Rokuonji" and "Kiyomizudera" in the tourist spot group. At this time, the features of "Rokuonji" will be gold color, gold leaf, glow, etc., the features of "Kiyomizudera" are the stage, the womb inside, the panoramic view etc. Because both are temples existing in Kyoto, features related to Kyoto and temples do not appear as unique features. Next, consider the case where "the Tokyo Metropolitan Government Building Observatories" and "Rokuonji" exist within the tourist spot group. Features of "Rokuonji" at this time will be Kinkakuji, a temple, golden color, Kyoto etc. Features of "the Tokyo Metropolitan Government Building Observatories" will be perspectives, night view, Shinjuku etc. If the categories of tourist spots are largely different, features as categories will appear. Also, it can show the features of the spot itself. In this research, when a certain spot compares with other spots in the set, we focus on the relative features that make it possible to clarify the features of each spot.

The spot differential vector is defined as formula 1. Is the value obtained by subtracting the average value of the spot vectors of the spots of the spot group excluding the spot for which the spot differential vector is found. $spot_{set} = \{s_1, s_2, \ldots, s_i, \ldots, s_n\}1$ is an already visited spot set or an unvisited spot set. s_i is a tourist spot in the set.

$$v_i = s_i - average(spot_{set} - s_i) \quad (1)$$

[1]https://github.com/neologd/mecab-ipadic-neologd/
[2]https://radimrehurek.com/gensim/models/doc2vec.html
[3]https://www.jalan.net/kankou/

C. Determination of explainable spot

From the feature difference vector v_i of the visited spot and the feature difference vector v_j of the unvisited spot, the relative feature similarity between the visited spot and the unvisited spot (Fig. 2). For the similarity calculation, use the cosine scale (formula 2).

$$cos(v_i, v_j) = \frac{v_{i1}v_{j1} + v_{i2}v_{j2} + \cdots + v_{in}v_{jn}}{\sqrt{v_{i1}^2 + \cdots + v_{in}^2} \times \sqrt{v_{j1}^2 + \cdots + v_{jn}^2}} \quad (2)$$

A correlation between the already visited spot having similarity of each feature vector of the already visited spot and each feature vector in the unvisited area of 0.125 or more and the highest similarity and the unvisited spot is performed. In addition, there are two ways to establish an association between an already visited spot and an unvisited spot.

1) how to relate to unvisited spots based on already visited spots
2) how to relate to already visited spots based on unvisited spots

In the method 1, if the spot in the unvisited area holds a plurality of features when using the already visited spot as the base, there are cases where the spot is associated with the same unvisited spot. In method 2, method 2 is used as it is possible to extract features of each spot from the already visited spot set based on the unvisited spot and associate with the unvisited spot. In addition, in this research, it is considered reasonable to base on unvisited spots to explain unvisited spots.

D. Extraction of explainable words for role

All tourist spots review words by using morphological analyzer "mecab-ipadic-NEologd". However, words obtained by using these words contain words that do not hold Japanese, and it is necessary to delete these noises. Specifically, delete particles, auxiliary verbs, rentaishi, symbols.

Since the explanation information of the visited spot and the unvisited spot associated in section III-C is presented to the user in word format, a review set of a certain tourist spot is assumed to be a document i and a spot where the word j for i appears When the number of occurrences of the set is $TF_{i,j}$, the word j is the number of documents in the spot set is DF_j, and the total number of spots in the spot set is $|D|$ The feature quantity of a word in a spot is defined by the formula 3.

$$word_{i,j} = TF_{i,j} \times IDF_j \quad (3)$$

$$IDF_j = log(\frac{|D|}{DF_j}) \quad (4)$$

In this method, for visited spots, the user inputs plurality of spots. By considering all reviews of each spot as one document at a time and by considering all reviews of the other spots as documents, the TFIDF value is calculated by the formula 3, 4, and use it as the feature words for each spot in the unvisited area.

Regarding the unvisited area, the user designates an area and inputs it. By considering all reviews of each spot in the area as one document and considering all reviews of the other spots as a document, the TFIDF value is calculated by the formula 3, 4, and use it as the feature words for each spot in the unvisited area.

The explanation information on the associated visited spots and unvisited spots associated with each other is determined by using harmonic averages according to feature words of each spot obtained by TFIDF. The harmonic mean is the reciprocal of the arithmetic mean of the reciprocal. Extracts commonly appearing words in the review document of the already visited spot and the review document of the unvisited spot. The score of the extracted word is defined by the formula 5. $word_{visited}$ and $word_{unvisited}$ indicate the TFIDF value of the visited spot and the TFIDF value of the unvisited spot, respectively. When the value of the word score is large, the TFIDF value of each of the visited spot and the unvisited spot is large, that is, the word has high importance in each document. Therefore, the top ten words of the word score are presented to the user as explanation information.

$$score = \frac{1}{\frac{1}{2}(\frac{1}{word_{visited}} + \frac{1}{word_{unvisited}})} \quad (5)$$

E. Example of explained unfamiliar tourist spots

IV. Evaluation Experiment

A. Experiment of feature word representing relation

B. Comparative experiment of correspondence

V. Conclusions and Future Work

Acknowledgment

The authors would like to thank...

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." Sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page, not here.

References

[1] T. Kurashima, T. Iwata, G. Irie and K. Fujimura., "Travel route recommendation using geotags in photo sharing sites", CIKM '10 Proceedings of the 19th ACM international conference on Information and knowledge management, pp.579-588, 2010

[2] R. Kitamura and T. Itoh, "Tourist Spot Recommmendation Applying Generic Object Recognition with Travel Photos", ITE Tech. Rep., Vol.42, No.12, AIT2018-94, pp.185-188, 2018

[3] A. J. Cheng, Y. Y. Chen, Y. T. Huang and Winston H. Hsu, "Personalized Travel Recommendation by Mining People Attributes from Community-Contributed Photos", MM '11 Proceedings of the 19th ACM international conference on Multimedia, pp.83-92, 2011

[4] K. J. Holyoak and P. Thagard, "Mental Leaps: Analogy in Creative Thought, MIT Press", Journal of Japanese Society for Artificial Intelligence, Vol.11, No.3, pp.489, 1996

[5] D. Gentner, "Structure-Mapping: A Theoretical Framework for Analogy", Cognitive Science, Vol.7, pp.155–170, 1983

[6] M. L. Gick and K. J. Holyoak, "Analogical Problem Solving", Cognitive Psychology, Vol.12, pp.306–355, 1980

[7] M. L. Gick and K. J. Holyoak, "Scheme Induction and Similarity in Analogical Transfer", Cognitive Psychology, Vol.15, pp.1–38, 1983

[8] Z. Chen and M. W. Daehler, "Positive and Negative Transfer in Analogical Problem-solving by 6-years-old Children", Cognitive Development, Vol.4, No.4, pp.327–344, 1989

[9] K. J. Holyoak and P. Thagard, "Analogical Mapping by Constraint Satisfaction", Cognitive Science, Vol.13, pp.295–355, 1989

[10] T. Kudo, K. Yamamoto and Y. Matsumoto, "Applying Conditional Random Fields to Japanese Morphological Analysis", Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237, 2004